Young scientists network
DAPHNE VISION

Scientists

Facilities

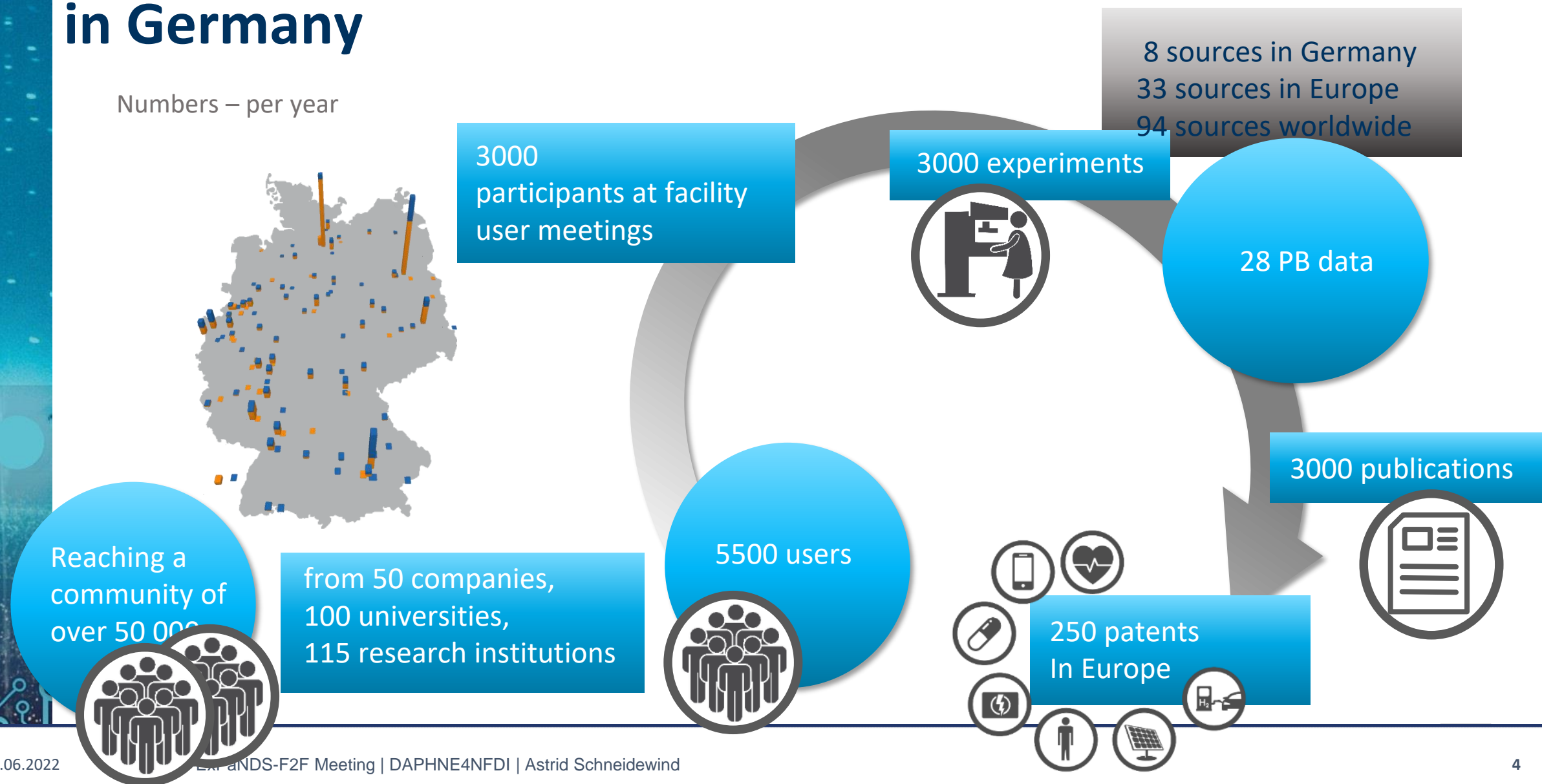# NFDI – German National Research Data Infrastructure

- Biodiversity, Ecology and Environmental Data
- Catalysis
- Chemistry
- Social, Educational, Behavioural and Economic Sciences
- Data in Plant Research
- Personal Health Data
- German Human Genome Phenome Archive
- Research data on tangible and intangible cultural assets
- Engineering Sciences

- Business, Economic and Related
- Data PHoton and Neutron Experiments
- Mathematical Research Data Initiative
- Condensed-Matter Physics and the Chemical Physics of Solids
- Data Science and Artificial Intelligence
- Materials Science and Engineering
- Particles, Universe, NuClei and Hadrons
- Language- and text-based Research Data

Cross-cutting: Base4NFDI
(proposed)

Research with Photons and Neutrons in Germany

Numbers – per year

8 sources in Germany
33 sources in Europe
94 sources worldwide

3000 participants at facility user meetings

3000 experiments

28 PB data

3000 publications

5500 users

Reaching a community of over 50 000

from 50 companies,
100 universities,
115 research institutions

250 patents
In Europe

DAPHNE4NFDI

Data from Photon and Neutron Experiments

Consortium

18 partners:
- University user groups
- Large scale facilities
- In addition: KFN + KFS

+ > 60 participants (without funding)

Kick-off meeting 1.12.2021

Task Area Leaders:

Anton Barty (DESY, Speaker)
Bridget Murphy (CAU)
Astrid Schneidewind (FZJ)
Christian Gutt (U Siegen)
Wiebke Lohstroh (TUM)

Sebastian Busch (Hereon)
Jan-Dierk Grunwald (KIT)
Frank Schreiber (U Tübingen)
Tobias Unruh (FAU)
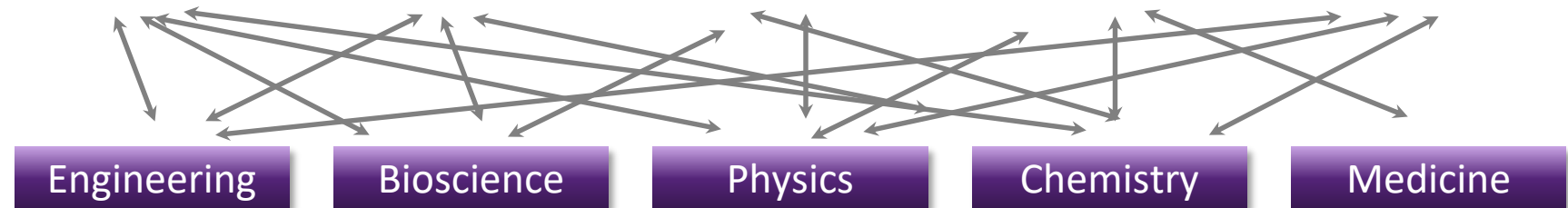
Research Community Impacts on Global Challenges

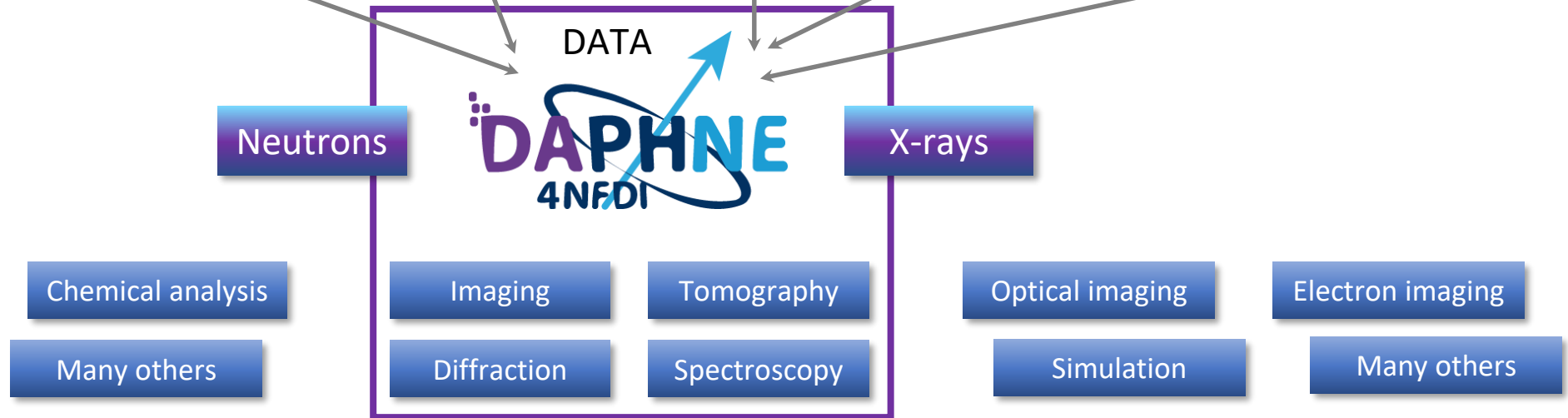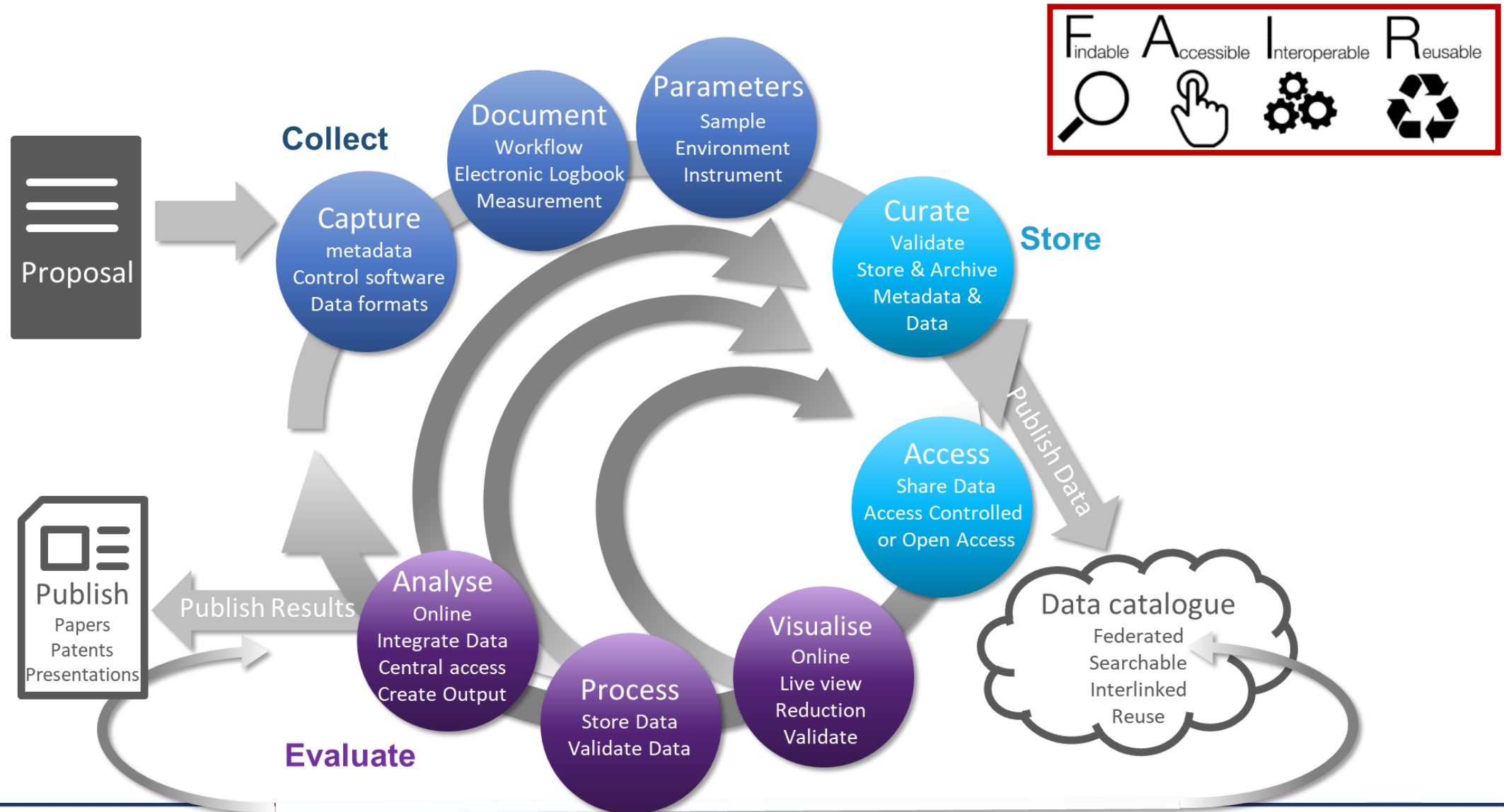**Impact extends far beyond the physics or materials science community**

Society challenges

Health	Environment	Energy	Materials	Transport	Knowledge

Research domains

Engineering	Bioscience	Physics	Chemistry	Medicine

DATA

Neutrons

X-rays

Analytic methods

Chemical analysis

Imaging

Tomography

Optical imaging

Electron imaging

Many others

Diffraction

Spectroscopy

Simulation

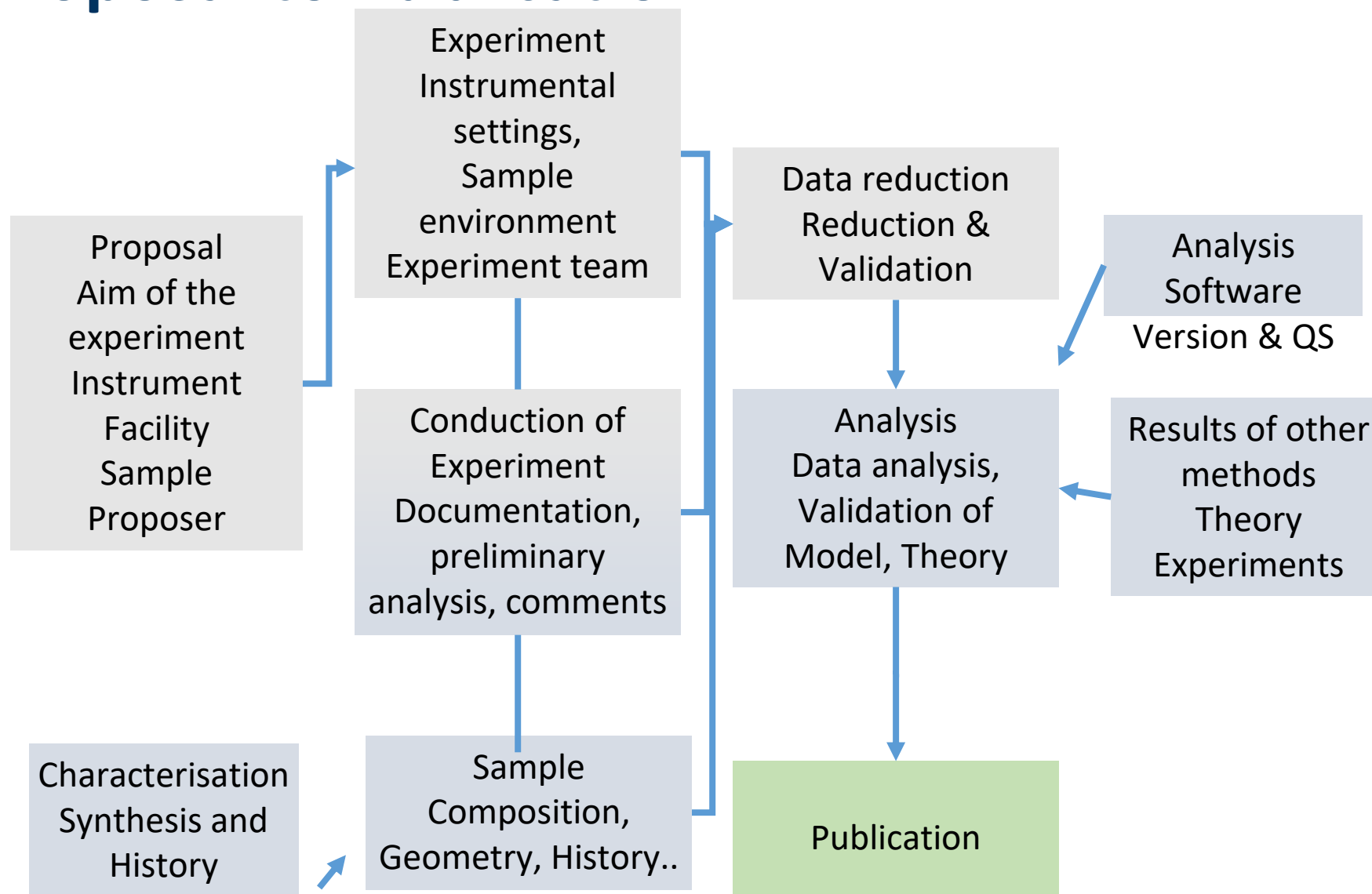Many others

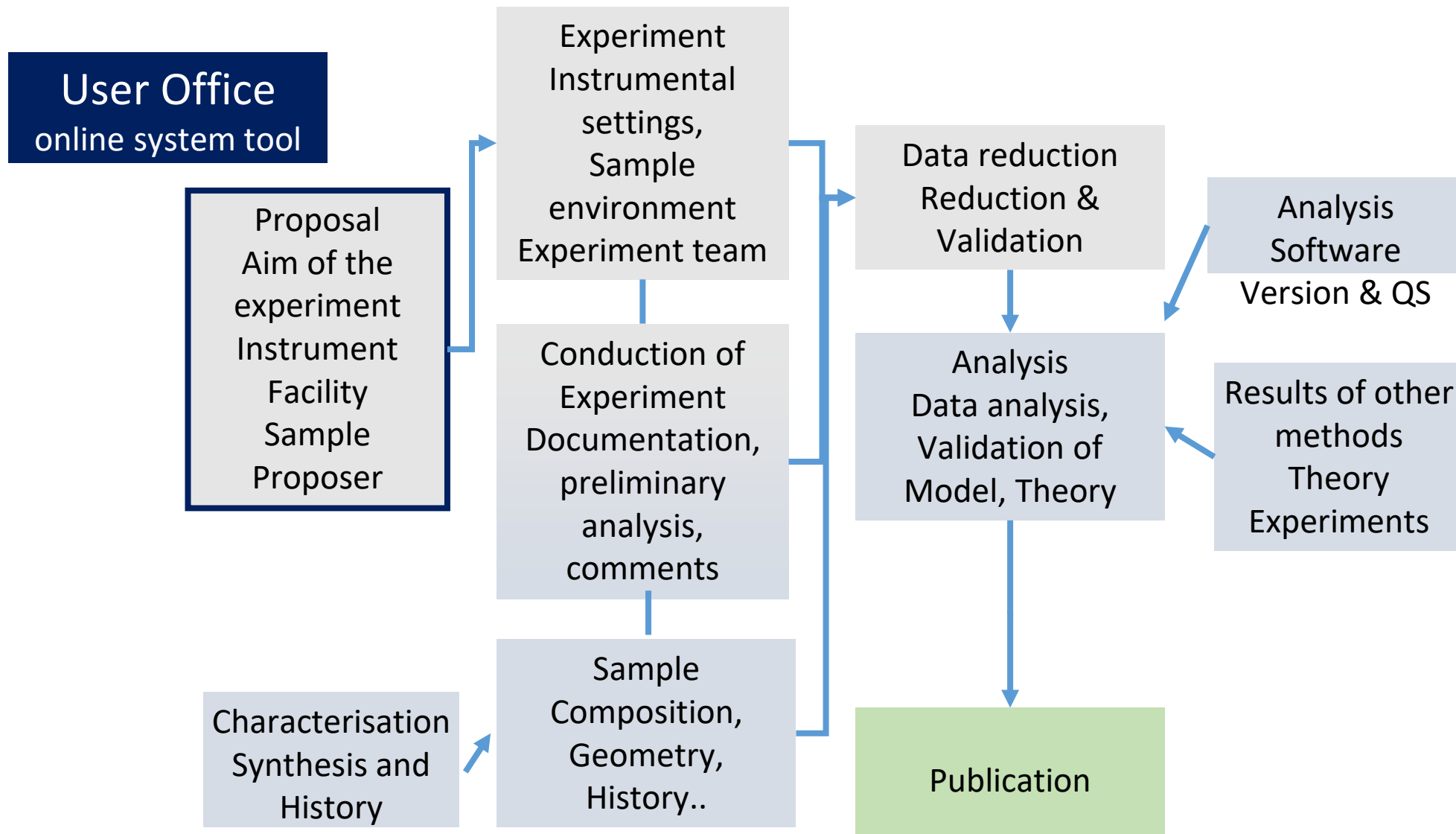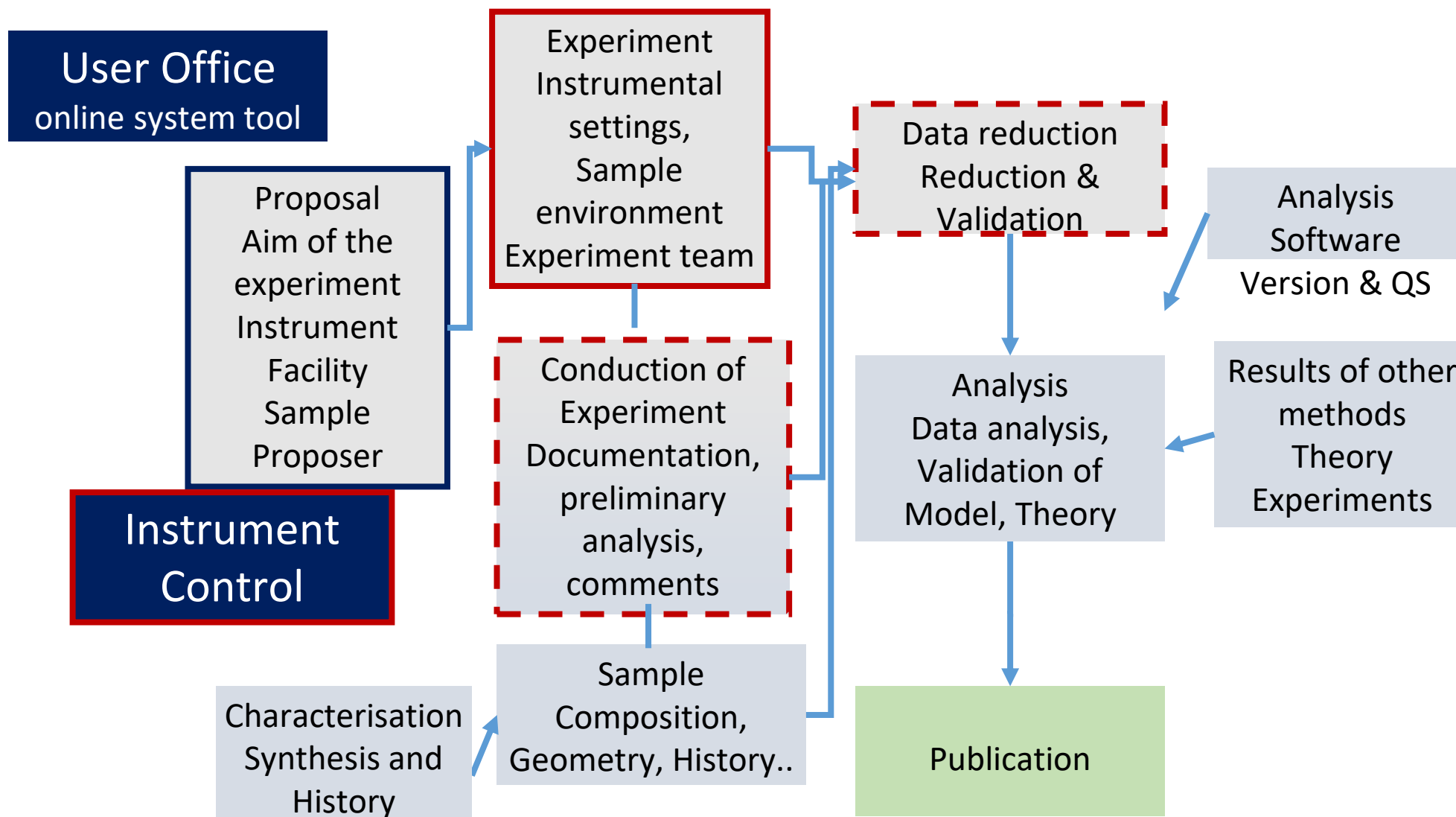# Research Data Management Strategy
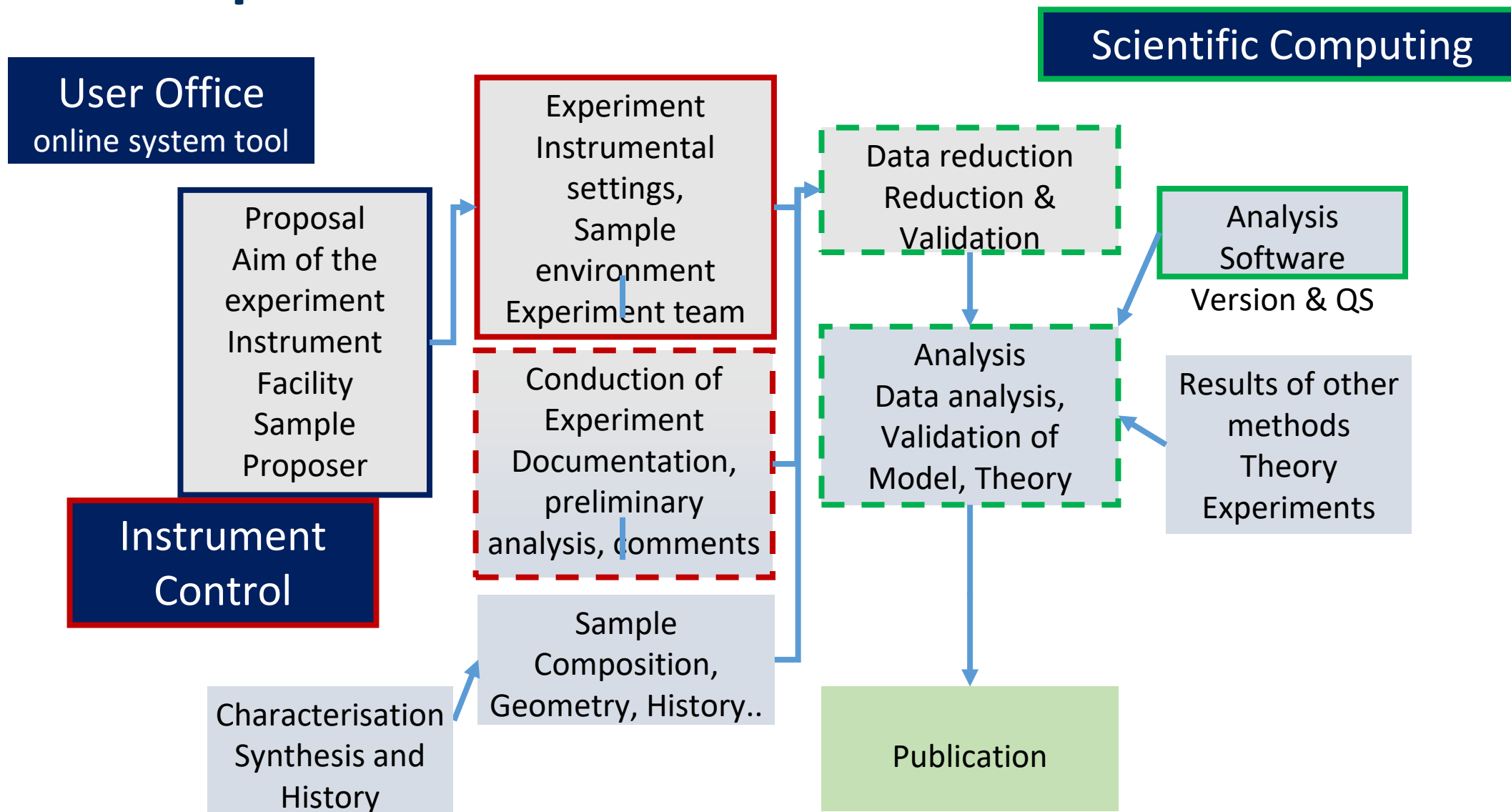
From Proposal to Publication

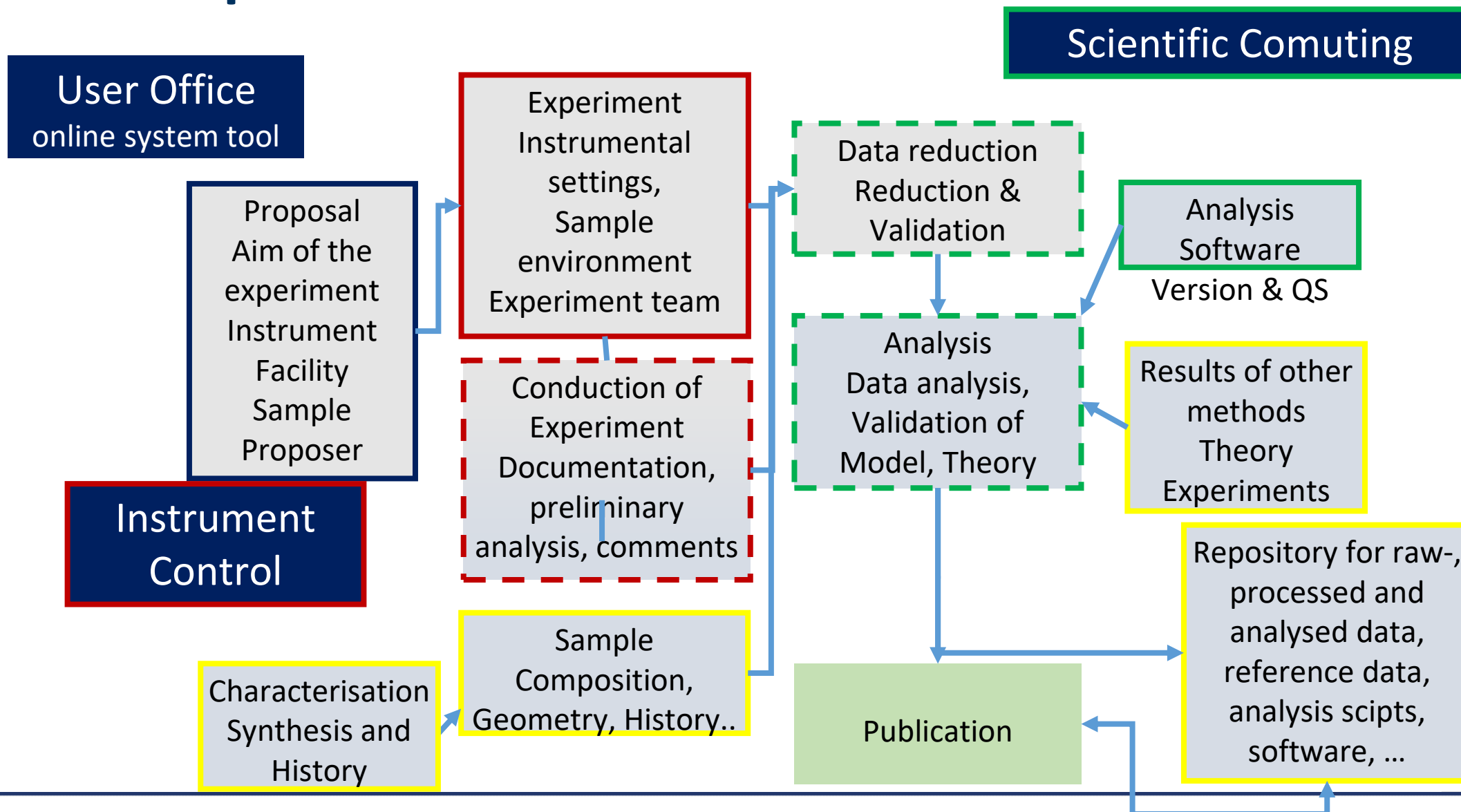# From Proposal to Publication

From Proposal to Publication

From Proposal to Publication

From Proposal to Publication

Scientific Comuting

User Office
online system tool

Proposal
Aim of the
experiment
Instrument
Facility
Sample
Proposer

Experiment
Instrumental
settings,
Sample
environment
Experiment team

Data reduction
Reduction &
Validation

Analysis
Software
Version & QS

Instrument
Control

Conduction of
Experiment
Documentation,
preliminary
analysis, comments

Analysis
Data analysis,
Validation of
Model, Theory

Results of other
methods
Theory
Experiments

Characterisation
Synthesis and
History

Sample
Composition,
Geometry, History..

Publication

Repository for raw-,
processed and
analysed data,
reference data,
analysis scipts,
software, …

DAPHNE4NFDI

TA1 : Managing data collection

Electronic laboratory notebooks (ELN)
High performance data formats
Sample Persistent identifiers (PID)

TA3 : Infrastructure for data and software reuse

Infrastructure for analysis software
Analysis software
ML methods

TA4 : Dissemination and outreach

Education & training
Workshops & use case

TA2 : (Meta)data repositories and catalogues

Data catalogues (and DOI)
Reference databases (RDB)
Catalogues for software

TA5 : External communication and policy
Policy
NFDI and European projects

TA6 : Management (DESY)

TA 1: Managing Data Collection

Capture (meta)data as early as possible in electronic form

- Specification of metadata standards – application to use domain specific use cases, (further) develop PaN ontologies

- Link to Sample PID – sample metadata schemata (e.g. IGSN): useroffices, STARS (LEAPS/LENS) facility-internal procedures

- Introduction of electronic Notebooks (ELN) („contextual description / annotated data'): tests

- Automatic capture of instrument / sample environment and sample data – conduction of the experiment: nicos for neutrons, SECoP (HMC)

- Aggregate in harmonized, self-explaining data format (NHDF5 / Nexus, OpenPMD)

ExPaNDS ontologies

Document Control Information	
Settings	**Value**
Document Identifier:	D3.2
Project Title:	ExPaNDS ontologies
Work Package:	WP3

https://zenodo.org/record/4806026#.YUh-kOdCRno

TA 2: (Meta)data Repositories and Catalogues

Searcharble, federated data repositories

- Specification of metadata standards – application to use domain specific use cases

- Establish federated, interlinked data repository for participating facilities: SciCat, SampleDB – workshop for university groups – policies and DOI

- Sample description and sample PID: STARS

- Repository for processed and analysed data: uses cases

- Reference data bases: use cases (EXAFS)

- Repository of analysis programs (e.g. https://software.pan-data.eu/)

- 'Search data' – not only publication-related

https://www.rcsb.org/

https://datacite.org

DataDOI tracks usage of open data

https://www.eosc-portal.eu
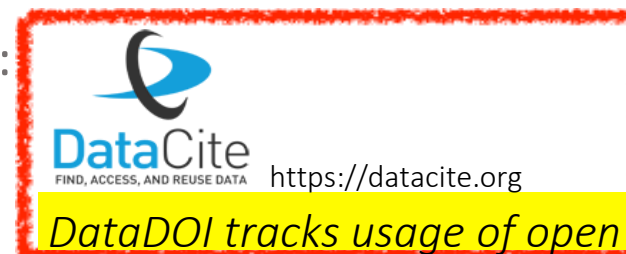
TA 3: Infrastructure for Data and Software Reuse

'FAIR software'

- Curate and deploy analysis software on cloud like infrastructures (currently frequently fragmented and diverse) for a number of use cases (tomography , EXAFS, small angle techniques, crystallography, TOF-diffraction, interface to ML methods)
- Automatized data analysis chains
- AAI infrastructure (NFDI?)
- Best practices for sustainable research software development
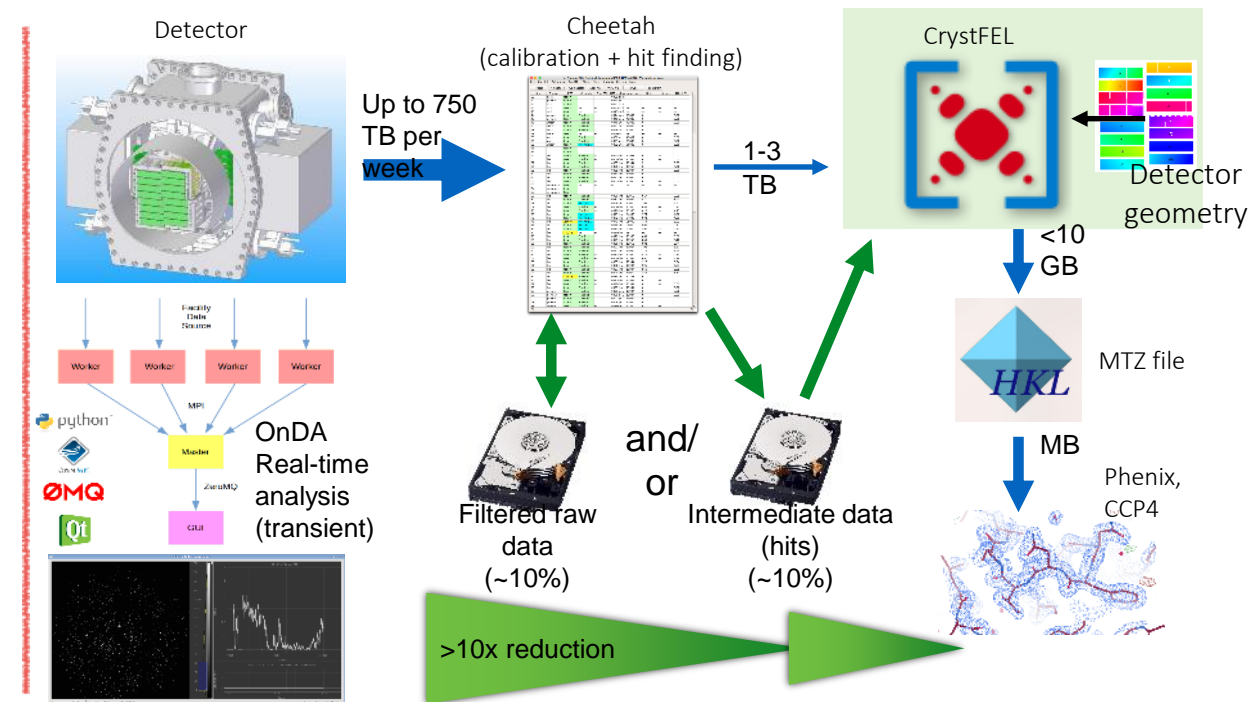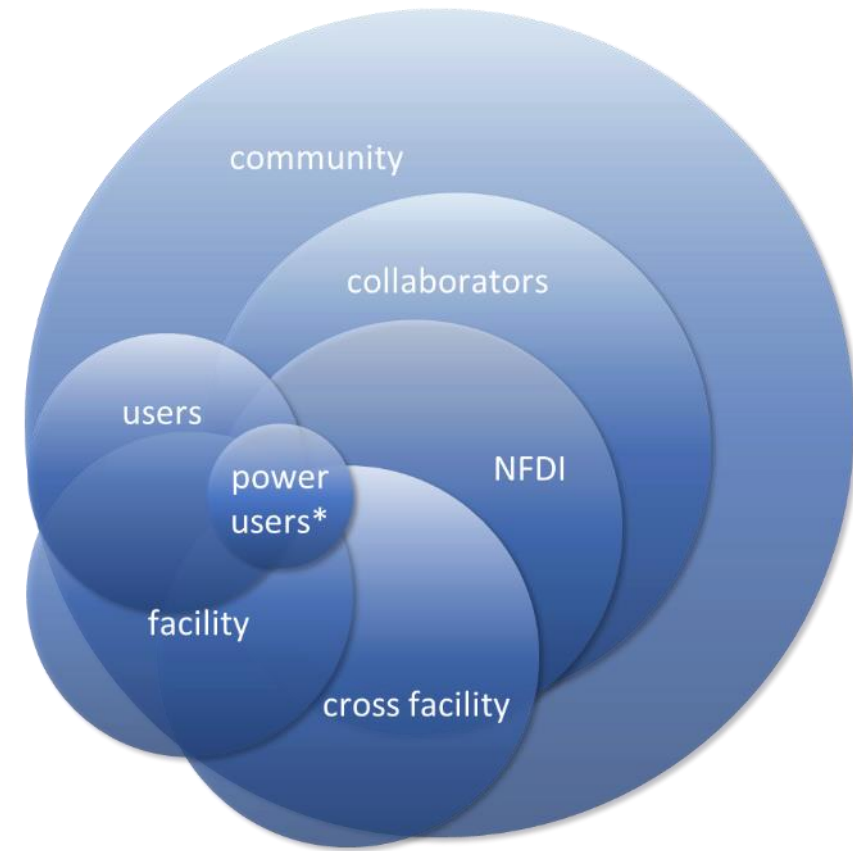
Task Area 4: Dissemination and Outreach

The NFDI consortium as a role model and educator

- Workshops and community building - (meta)data definition and ontologies

- Discuss and Support use cases

- Young researches network – **DAPHNE VISION**

- Data managment as part of the curriculum

- Connect society and industry to highlight developments

- Encourage to re-use

Task area 5: External Communication and Policy

Integration into the international research community

- DAPHNE is embedded in a worldwide network of > 30.000 synchrotron and neutron users
- Cross-consortia activities – interfaces and interlinkages to other NFDI consortia
- Organizational structures exists: European user organizations and facility organization
- Connects to European open science cloud X-ray and neutron data projects PaNOSC and ExPaNDS
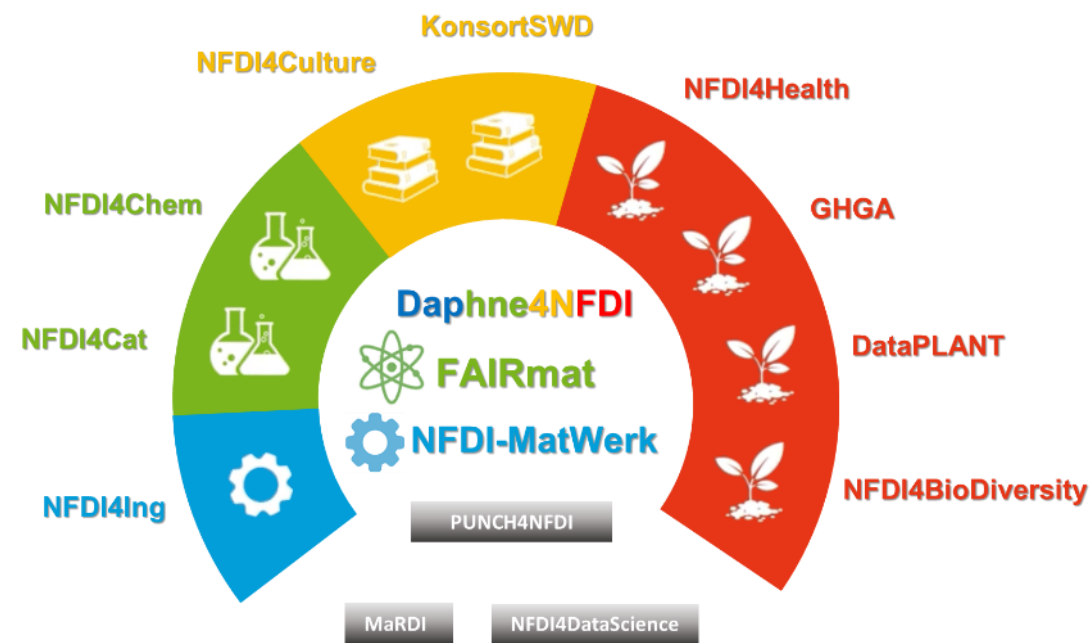
Use Cases – DAPHNE as a Role Model and Educator

Biomaterials
x-ray imaging
LMU – Uni Göttingen

Energy and battery materials, catalysis
Tomography
TUM – MLZ – BAM – hereon – HZB - KIT

Amorphous materials for catalysis
x-ray absorption spectroscopy
KIT – TUB – Uni Wuppertal

Soft matter and liquid interfaces
x-ray reflectivity
Uni Kiel – Uni Tübingen

Proteins & Food science
Diffraction (small and wide angle)
Spectroscopy
FAU– Uni Tübingen – EMBL - Uni Kiel

Magnetic structures
Ultrafast / Magnetic x-ray scattering
DESY – Uni Siegen

Dynamics
Correlation spectroscopy - XPCS
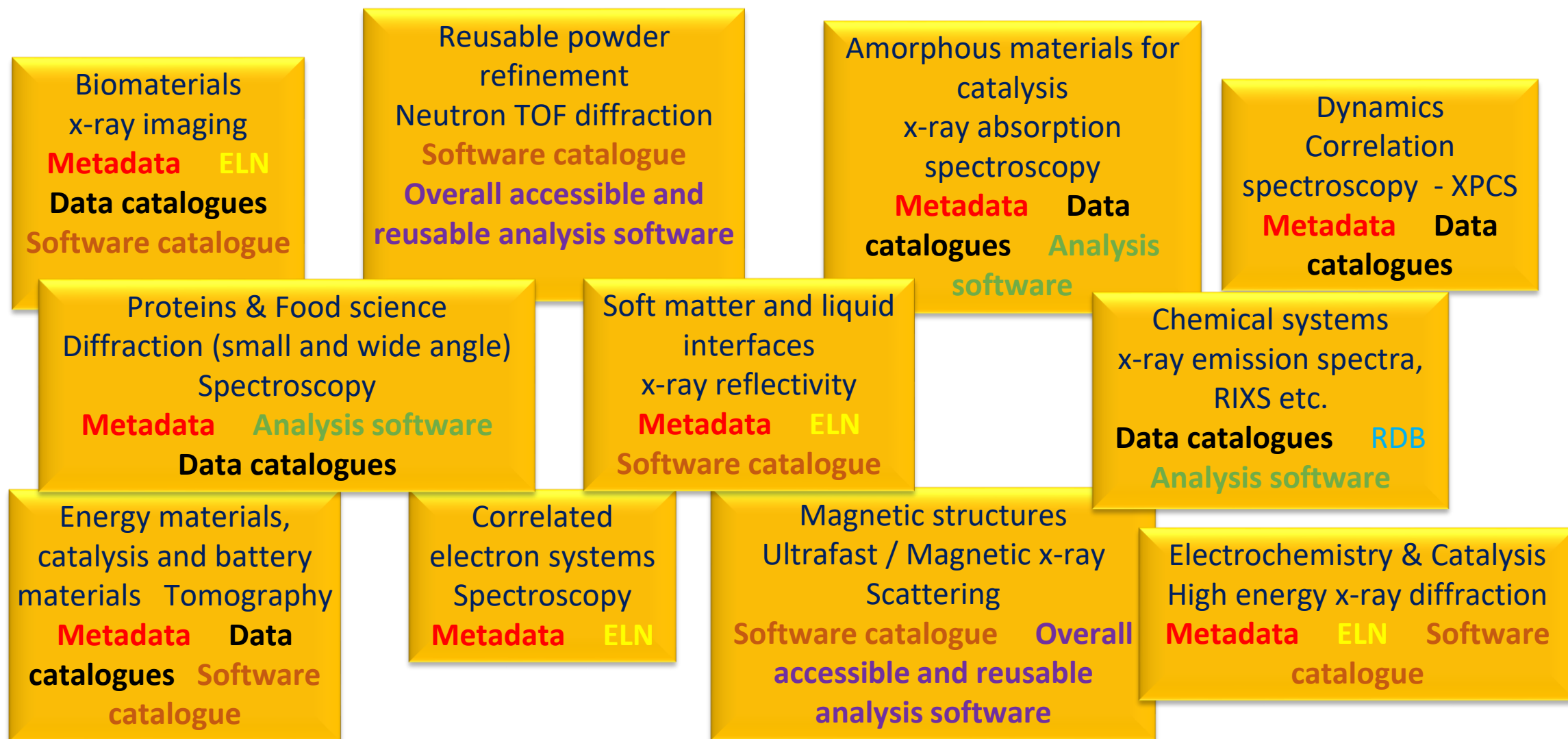Uni Siegen - EuXFEL

Correlated electron systems
Spectroscopy
KIT – FZJ - MLZ

Electrochemistry & Catalysis
High energy x-ray diffraction
HZDR – Uni Kiel - DESY

Reusable powder refinement
Neutron TOF diffraction
FZJ – MLZ – ESS - RWTH

Chemical systems
x-ray emission spectra, RIXS etc.
KIT – ESRF - DESY

Use Cases – DAPHNE as a Role Model and Educator

**Biomaterials
x-ray imaging**
Metadata ELN
Data catalogues
Software catalogue

**Reusable powder
refinement
Neutron TOF diffraction**
Software catalogue
Overall accessible and
reusable analysis software

**Amorphous materials for
catalysis
x-ray absorption
spectroscopy**
Metadata Data
catalogues Analysis
software

**Dynamics
Correlation
spectroscopy - XPCS**
Metadata Data
catalogues

**Proteins & Food science
Diffraction (small and wide angle)
Spectroscopy**
Metadata Analysis software
Data catalogues

**Soft matter and liquid
interfaces
x-ray reflectivity**
Metadata ELN
Software catalogue

**Chemical systems
x-ray emission spectra,
RIXS etc.**
Data catalogues RDB
Analysis software

**Energy materials,
catalysis and battery
materials Tomography**
Metadata Data
catalogues Software
catalogue

**Correlated
electron systems
Spectroscopy**
Metadata ELN

**Magnetic structures
Ultrafast / Magnetic x-ray
Scattering**
Software catalogue Overall
accessible and reusable
analysis software

**Electrochemistry & Catalysis
High energy x-ray diffraction**
Metadata ELN Software
catalogue

DAPHNE

Multi-dimensional matrix:

Scientific areas – Scattering technique – Institutions – Data related topics

Large overlap between use cases in one/some of these dimensions
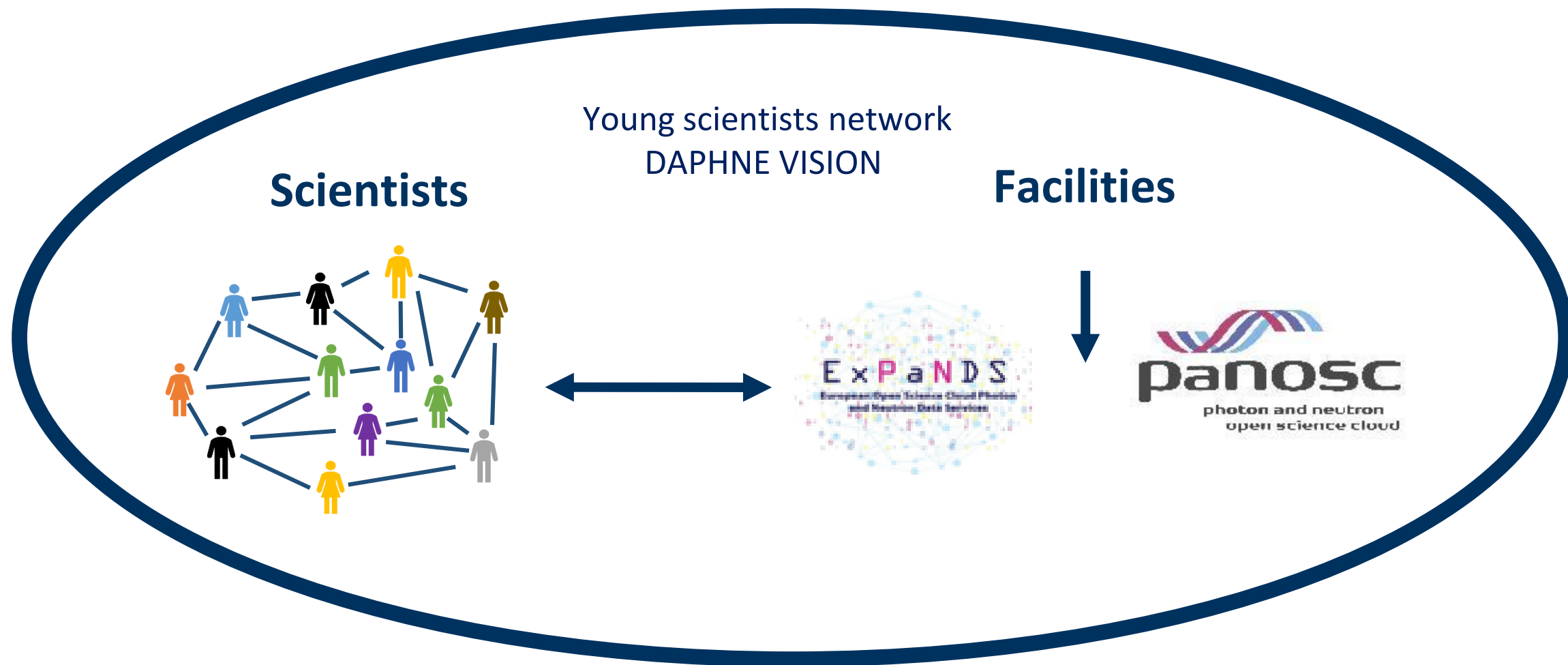
Communication! Networking!

DAPHNE: Mostly Scientists
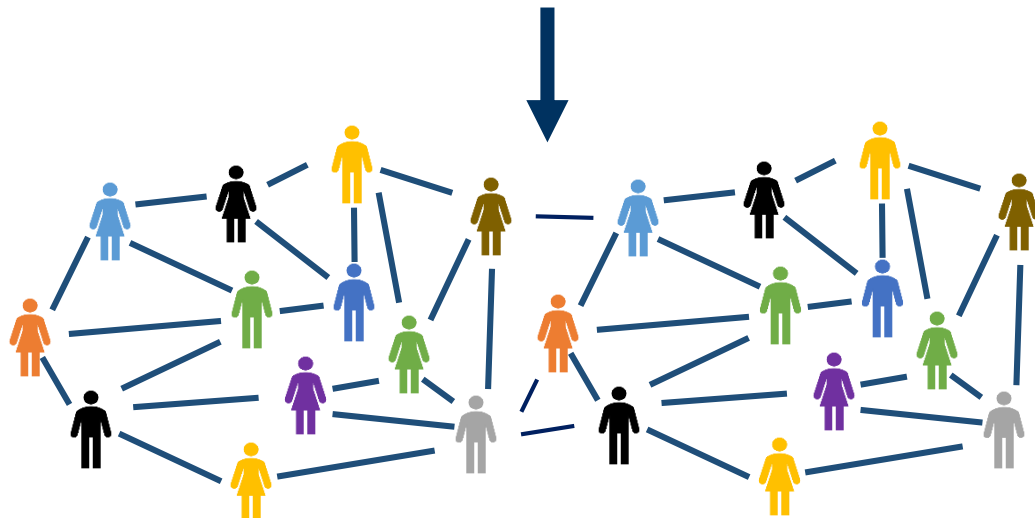
Facilities: MLZ, DESY, HZDR, HZB
ESRF, ILL ESS, EuXFEL

DAta for PHoton and Neutron Experiments

https://www.daphne4nfdi.de/

a.schneidewind@fz-juelich.de
contact@daphne-nfdi.de

Thank you for your attention!